



Gemeente
Amsterdam

How to use *our own* spatial network to perform graph analytics



Nico de Graaff
Spatial Data Scientist
Onderzoek, Informatie & Statistiek



Met dank aan:



Johan Olsthoorn (VNOR)

Martijn Kobus (VNOR)

Manilde van der Oord (OIS)

Martijn de Jong (OIS)

Niels Prins (voormalig inhuur)

DataTeam Ruimte (OIS)



Waarom netwerkanalyse?

Perceptie van afstand

- hemelsbreed (euclidisch, spherisch, ellipsoidisch)
- Manhattan distance (city block)
- Chebyshev Distance (Chessboard)
- Cosine distance



Waarom netwerkanalyse?

afstand over een netwerk

- ‘cost’ om van A naar B (of verder naar C) te komen
- kortste pad versus het meest optimale pad, travelling salesman
- locatieplanning, serviceniveau’s
- weight Matrix voor spatial autocorrelation , spatial regression



Netwerkanalyse: wat is er voor nodig?

- technologie, tools
- algoritmes
- Data!!
- ...en skills..

X
X
X

Spark GraphX
R iGraph QGIS
OSMnx Pandana
NetworkX
PgRouting
python-packages
PySal SpatialAccess
ArcGIS NetworkAnalyst
GIS-tooling
UrbanSim
NEO4J

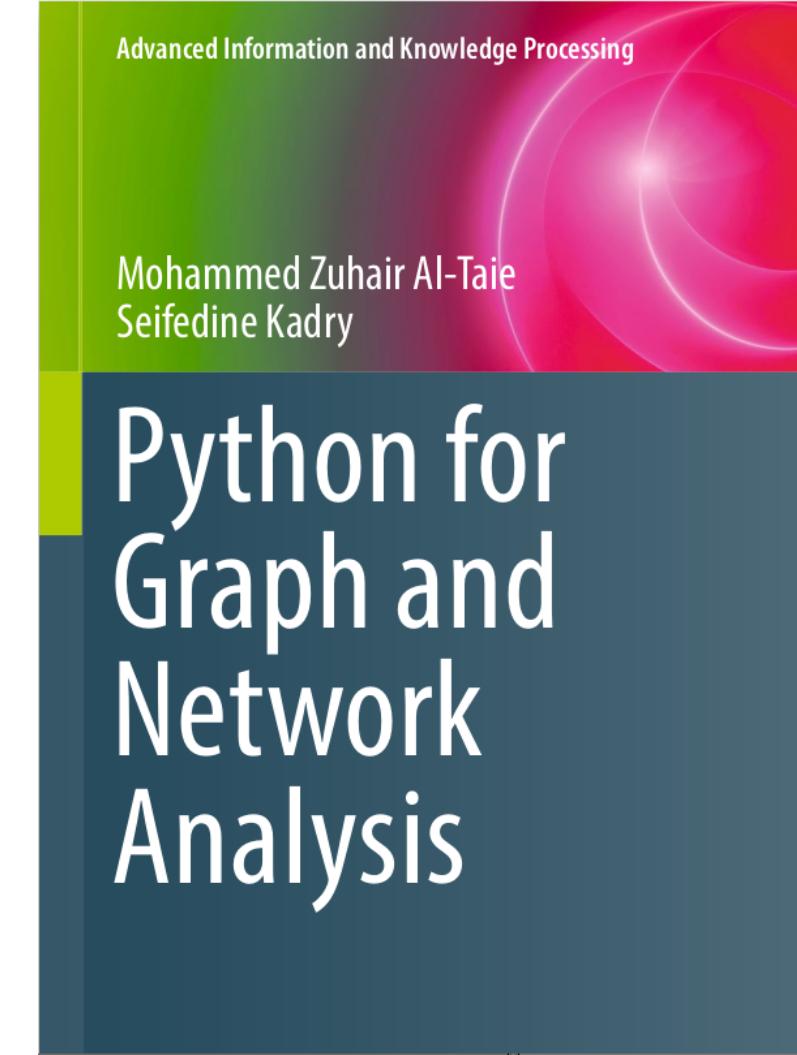
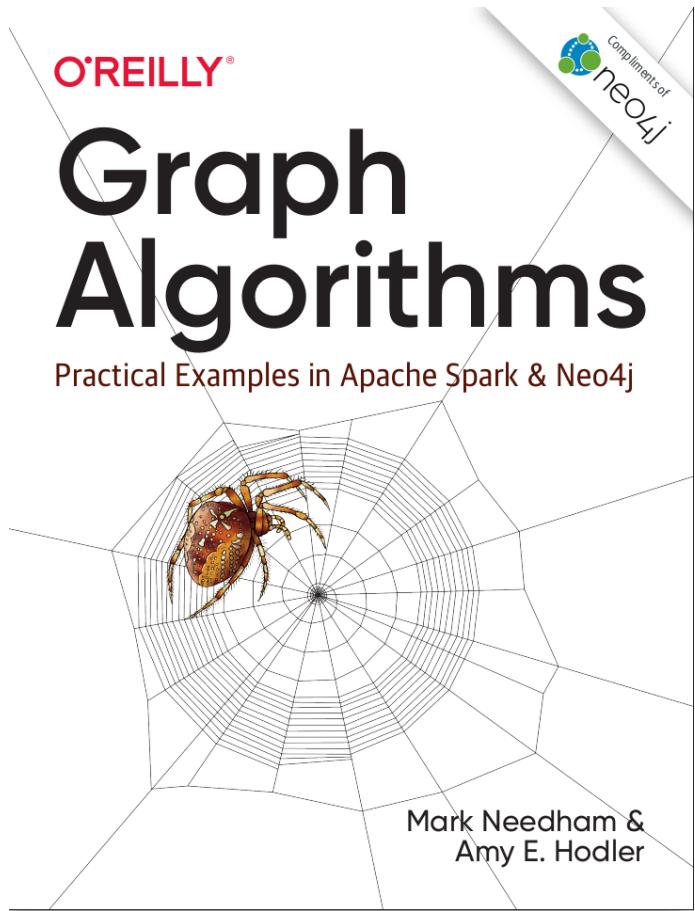


Algo's

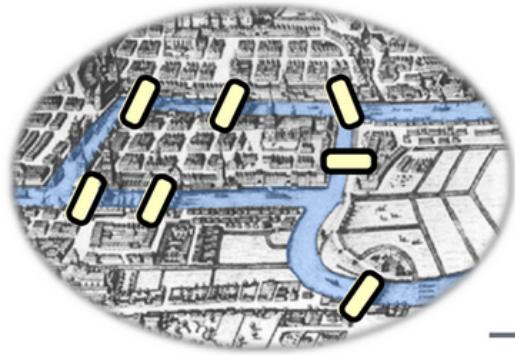
Algorithm type	What it does	Example use
Breadth First Search	Traverses a tree structure by fanning out to explore the nearest neighbors and then their sublevel neighbors	Locating neighbor nodes in GPS systems to identify nearby places of interest
Depth First Search	Traverses a tree structure by exploring as far as possible down each branch before backtracking	Discovering an optimal solution path in gaming simulations with hierarchical choices
Shortest Path Variations: A*, Yen's	Calculates the shortest path between a pair of nodes	Finding driving directions between two locations
All Pairs Shortest Path	Calculates the shortest path between <i>all pairs of nodes</i> in the graph	Evaluating alternate routes around a traffic jam
Single Source Shortest Path	Calculates the shortest path between a <i>single root node</i> and <i>all other nodes</i>	Least cost routing of phone calls
Minimum Spanning Tree	Calculates the path in a connected tree structure with the smallest cost for visiting all nodes	Optimizing connected routing, such as laying cable or garbage collection
Random Walk	Returns a list of nodes along a path of specified size by randomly choosing relationships to traverse.	Augmenting training for machine learning or data for graph algorithms.

XXX

Literatuur -> skills

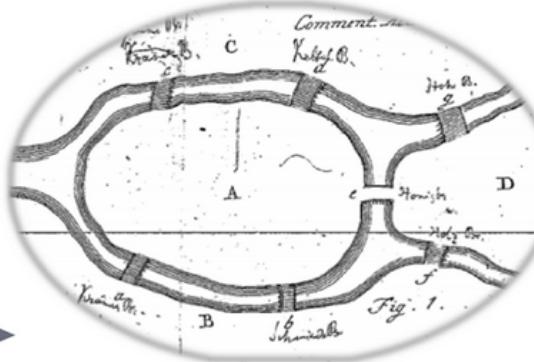


X intermezzo: wat is een netwerk? X (of graph)



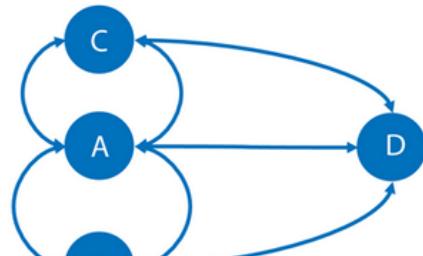
Walking the Bridges of Königsberg

4 Main areas of Königsberg with 7 Bridges.
Can you cross each bridge only once
and return to your starting point?



Euler's Insight

The only relevant data is the main areas
and the bridges *connecting* them.



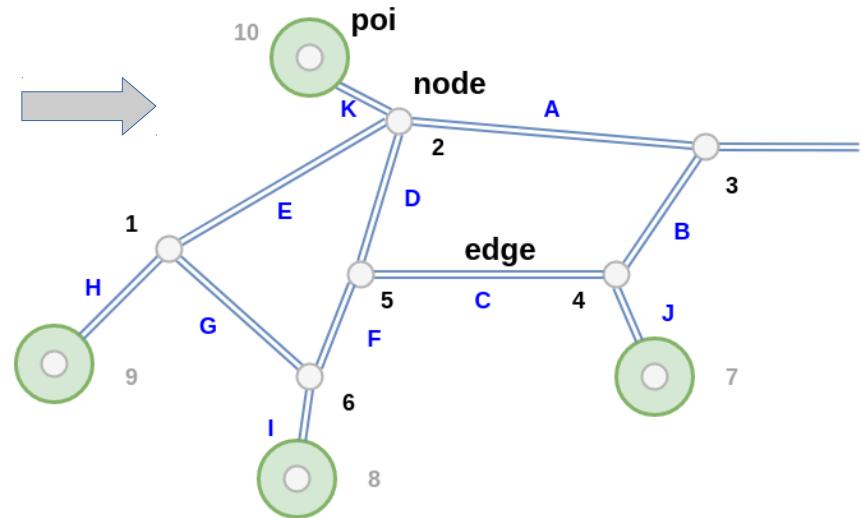
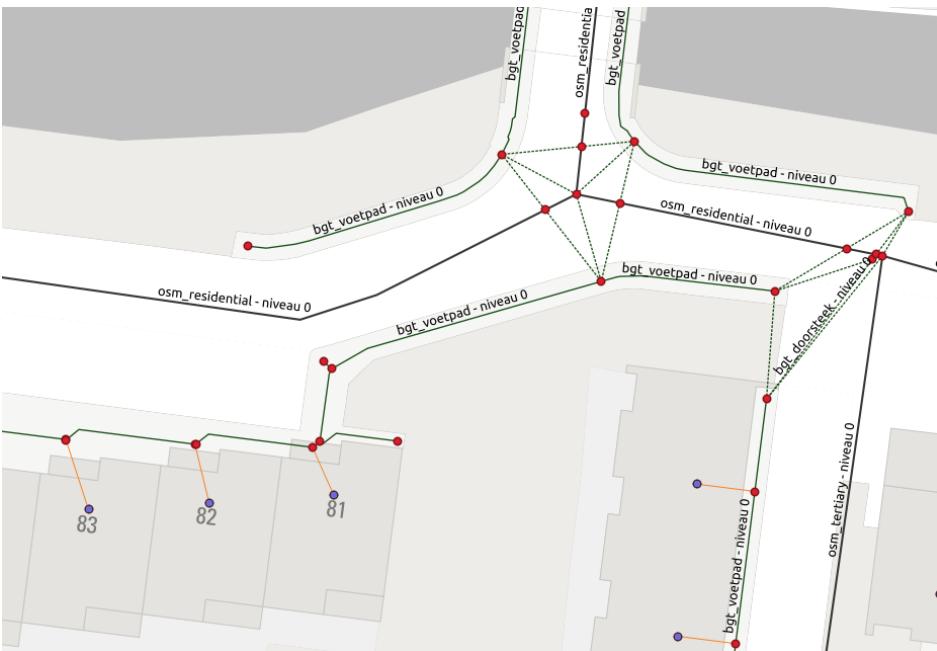
Origins of Graph Theory

Euler abstracted the problem and created
generalized rules based on nodes and
relationships that apply to any
connected system.

Figure 1-1. The origins of graph theory. The city of Königsberg included two large islands connected to each other and the two mainland portions of the city by seven bridges. The puzzle was to create a walk through the city, crossing each bridge once and only once.

X
X
X

Van aslijnen naar topologie (Graph)



edge_id	edge_weight	node_startpoint	node_endpoint
A	121.50	2	3
B	85.43	3	4
C	118.12	4	5
F	64.23	5	6
D	82.14	5	2



Waarom een eigen netwerk?

- grip op eigenschappen van het netwerk i.c.m. algorithmes
- berekenen van een groot aantal routes (distancematrix)
 - loopafstanden woningen - scholen: c.a. **50 miljoen** routes
 - loopafstanden woningen - afvalcontainerlocaties: c.a. **1.2 miljoen (dagelijks!)**
- flexibele toepasbaarheid met verschillende technologie
- kosten (licentie Google api)

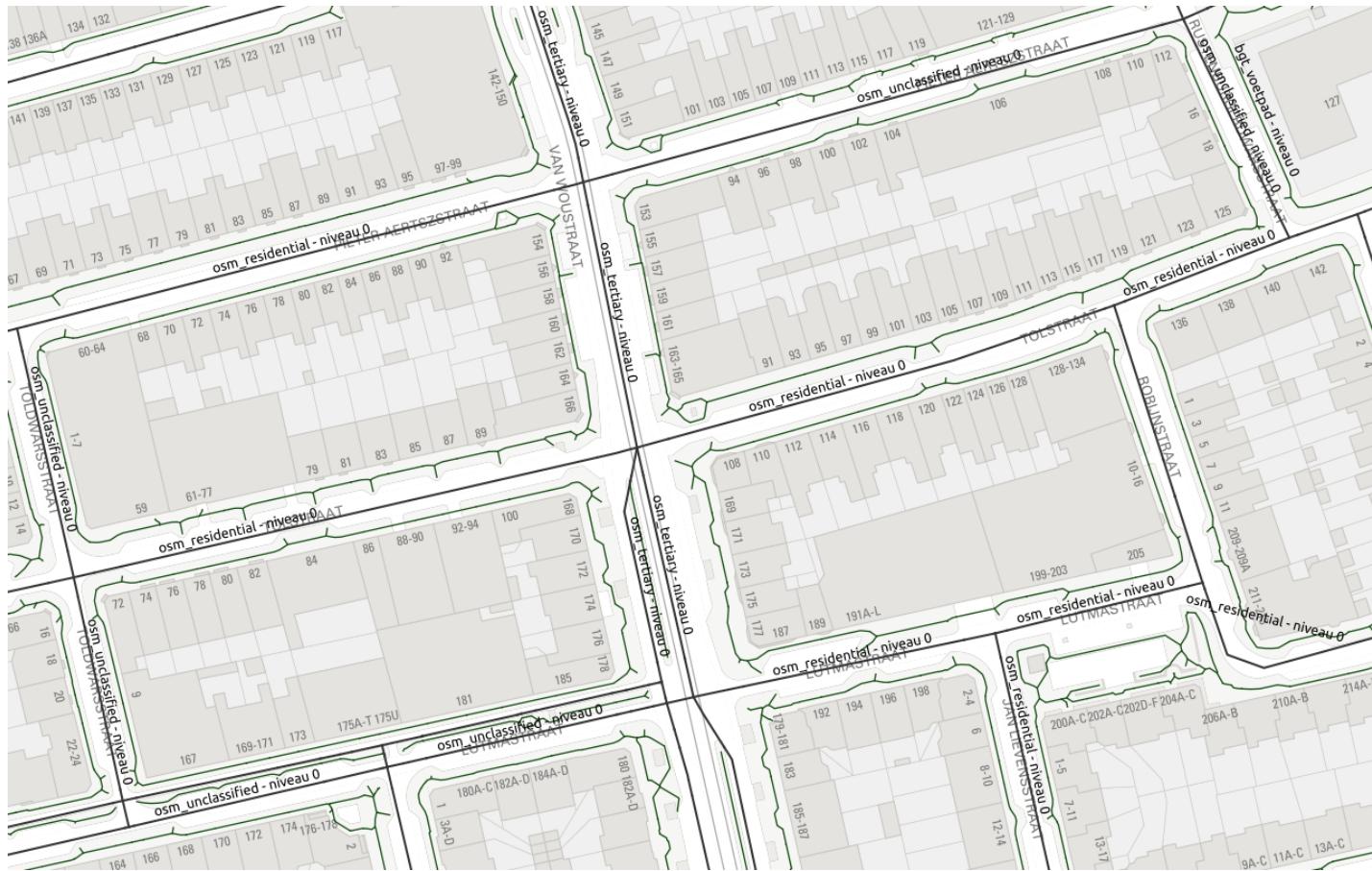


Fasering opbouw: OSM



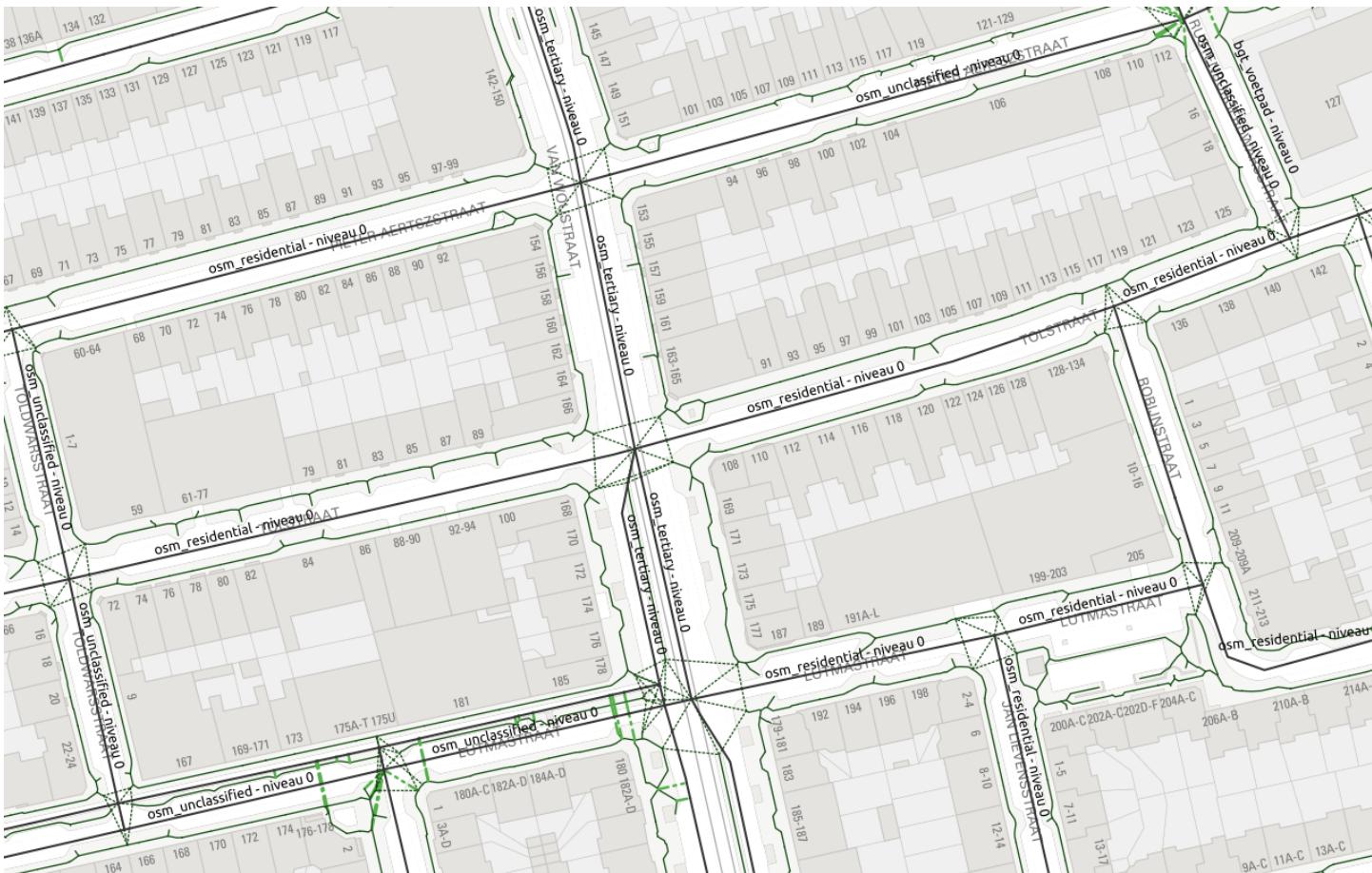
X
X
X

Fasering opbouw: OSM+BGТ-voetpad



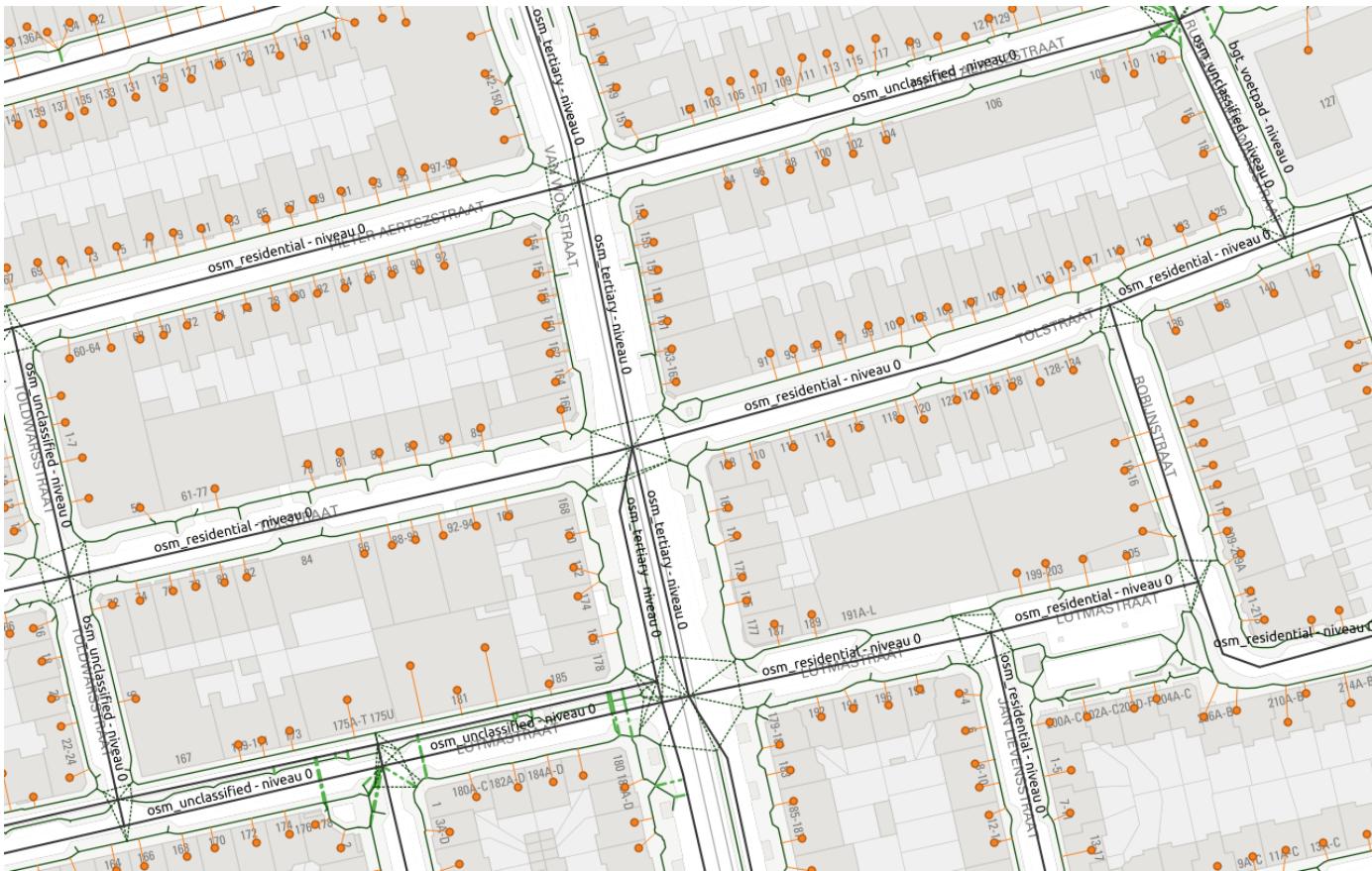
X
X
X

Fasering opbouw: OSM+BGD+oversteek





Fasering opbouw: OSM+BGT+oversteek+POI



XXX

Dataductie: clusteren 'adressen'



XXX

Dataductie: clusteren 'adressen'





Zelf poi's toevoegen

- The easy way: join met adressen
- the convenient way: nearest neighbour node
- the hard way: connect poi to roads

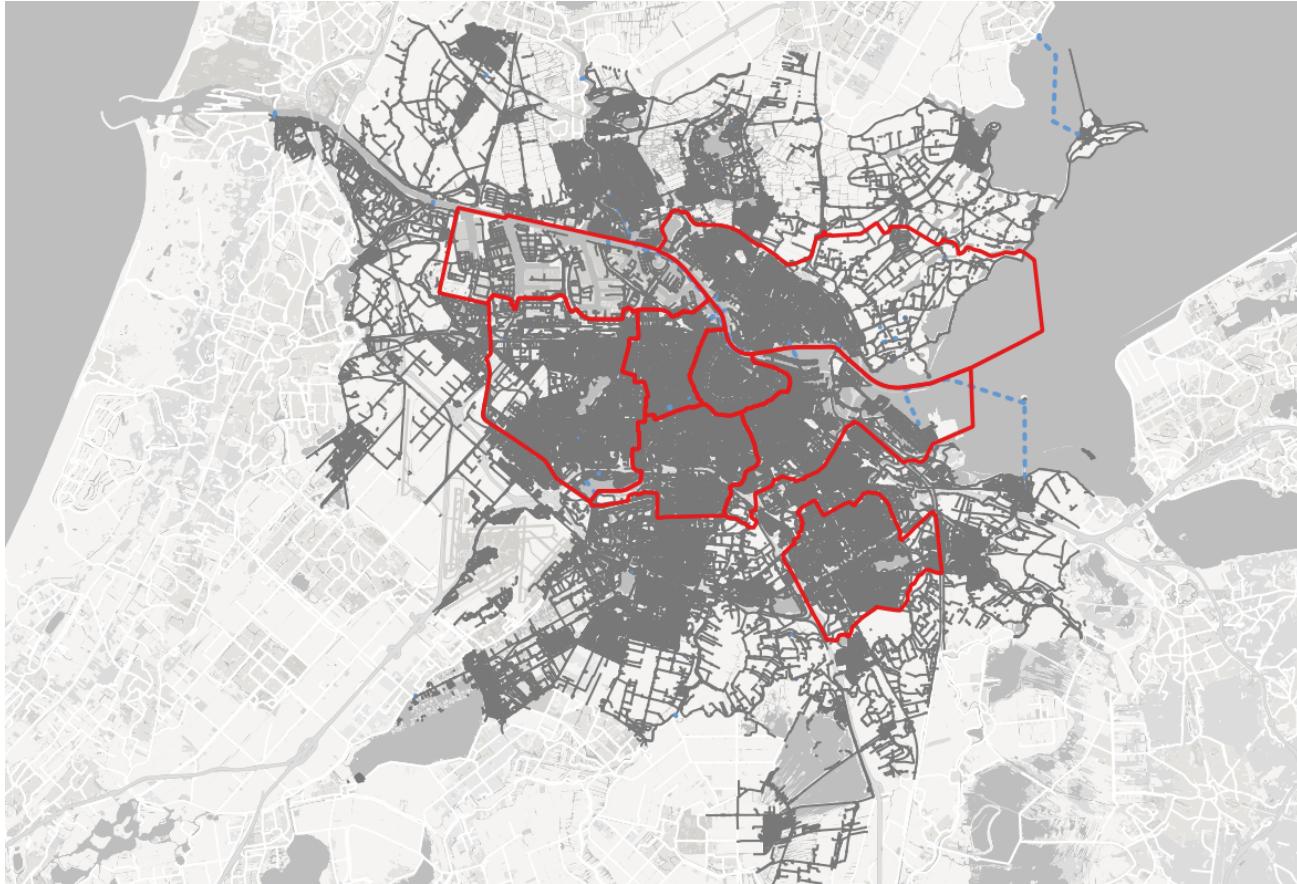


Aandachtspunten

1. aanwezigheid voetpaden in sub-urbaan en ruraal gebied beperkt
2. actualiteit osm en bgt, bij hoge eisen aslijnen handmatig toevoegen
3. persistente opslag Graph voor ad-hoc analyses en ontsluiting
4. Is OSM is nog nodig? voor veerverbindingen en max. snelheid
(rijbaan uit BGT, TOP10NL en NDW verkeersborden ?)
5. uitbreiden POI: OV-haltes, park, sportterrein

XXX

Gebruik





dataset: data.amsterdam.nl

Netwerk in bestandsvorm: GeoPackage met drie 'layers'

- edges: **adam_netwerk_voetfiets_edges**
- nodes en poi: **adam_netwerk_voetfiets_nodes_poi**
- adressen (verblijfsobject, ligplaats, standplaats):
adam_netwerk_voetfiets_poi_adressen



Dank voor de aandacht!